**Group Members:**

*MahnoorAamir 211980063*

*AyeshaBibi 211980004*

**Course: Data Visualization**

**Assignment#1**

**Report on Rainfall Data Visualization in Pakistan (1901-2016)**

**Introduction:**

This report presents an analysis of the monthly and annual rainfall data in Pakistan from 1901 to 2016. The data is visualized through various plots to understand trends, seasonal variations, and distribution patterns. The analysis is performed using R, utilizing the ggplot2 and dplyr packages for data manipulation and visualization.

**Data Loading and Preparation:**

**Libraries and Data Loading:**

First, we load the necessary libraries and the dataset:

# Load the libraries

library(ggplot2)

library(dplyr)

# Load the dataset

rainfall\_data <- read.csv("rainfall\_1901\_2016\_pak.csv")

# View the first few rows of the dataset

head(rainfall\_data)

View(rainfall\_data)

Here, we load the ggplot2 and dplyr libraries. The dataset is read from a CSV file named "rainfall\_1901\_2016\_pak.csv". The head function is used to display the first few rows, and View opens the dataset in a viewer for inspection.

**Data Transformation:**

# Convert Month to a factor with proper ordering

rainfall\_data$Month <- factor(rainfall\_data$Month, levels = c("January", "February", "March", "April", "May", "June", "July", "August", "September", "October", "November", "December"))

# Create a Date column for better plotting

rainfall\_data$Date <- as.Date(with(rainfall\_data, paste(Year, Month, "1", sep = "-")), "%Y-%B-%d")

 The Month column is converted to a factor with a specified order to ensure correct chronological plotting.

 A new Date column is created by combining the Year and Month columns, setting the day to the first of each month. This helps in plotting time series data.

## Data Visualization:

### Monthly Rainfall Over Time:

We plot the monthly rainfall data over time:

# Plot monthly rainfall over time

ggplot(rainfall\_data, aes(x = Date, y = Rainfall....MM.)) +

geom\_line(color = "blue") +

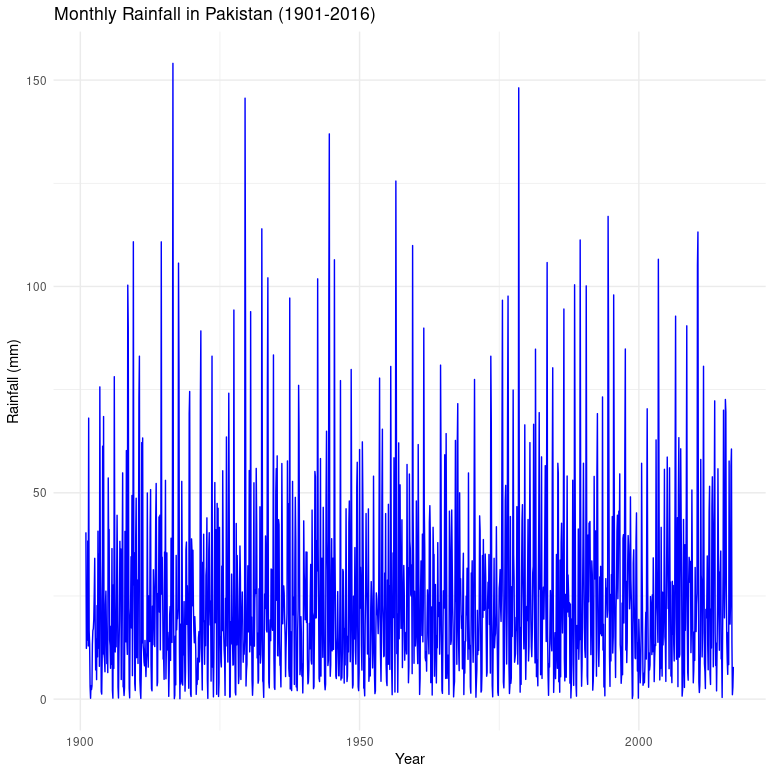
labs(title = "Monthly Rainfall in Pakistan (1901-2016)",

x = "Year",

y = "Rainfall (mm)") +

theme\_minimal()

This line plot visualizes monthly rainfall over the entire period, showing trends and variations over time.



### Annual Rainfall Over Time:

We calculate and plot annual rainfall:

# Calculate annual rainfall

annual\_rainfall <- rainfall\_data %>%

group\_by(Year) %>%

summarise(Total\_Rainfall = sum(Rainfall....MM.))

# Plot annual rainfall over time

ggplot(annual\_rainfall, aes(x = Year, y = Total\_Rainfall)) +

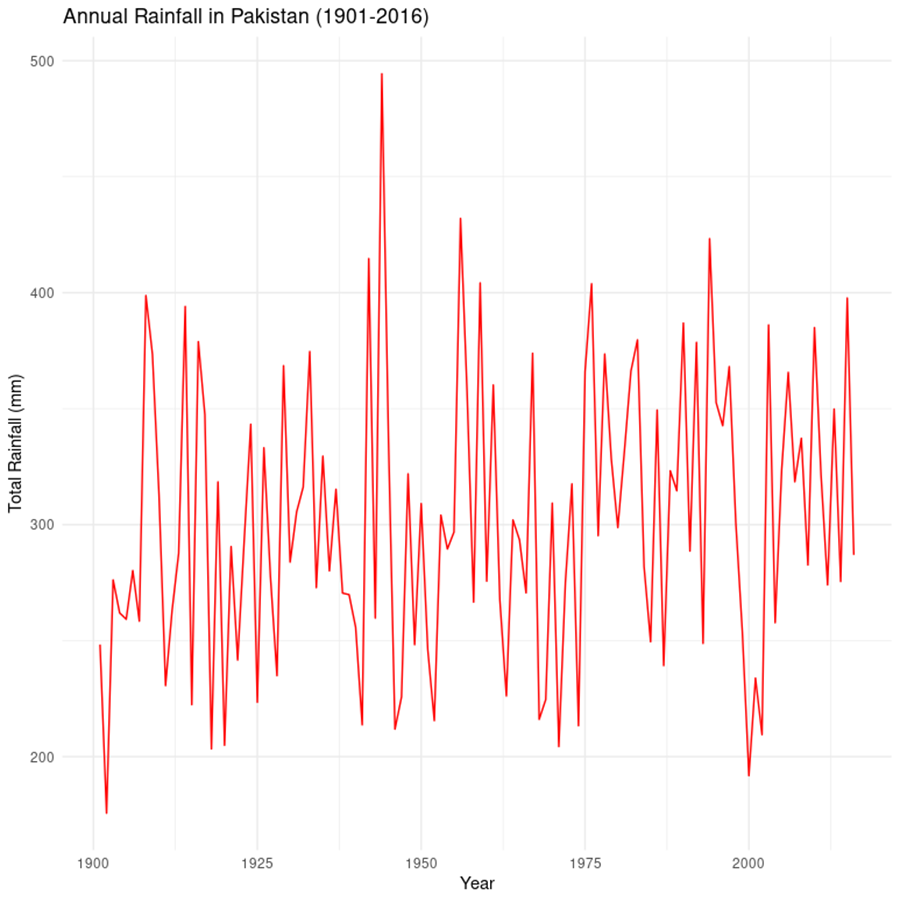
geom\_line(color = "red") +

labs(title = "Annual Rainfall in Pakistan (1901-2016)",

x = "Year",

y = "Total Rainfall (mm)") +

theme\_minimal()



• Annual rainfall is calculated by summing the monthly rainfall for each year.

• This line plot shows the total annual rainfall, highlighting long-term trends and potential anomalies.

### Monthly Rainfall Distribution:

We create a boxplot to examine the distribution of monthly rainfall:

# Boxplot of monthly rainfall

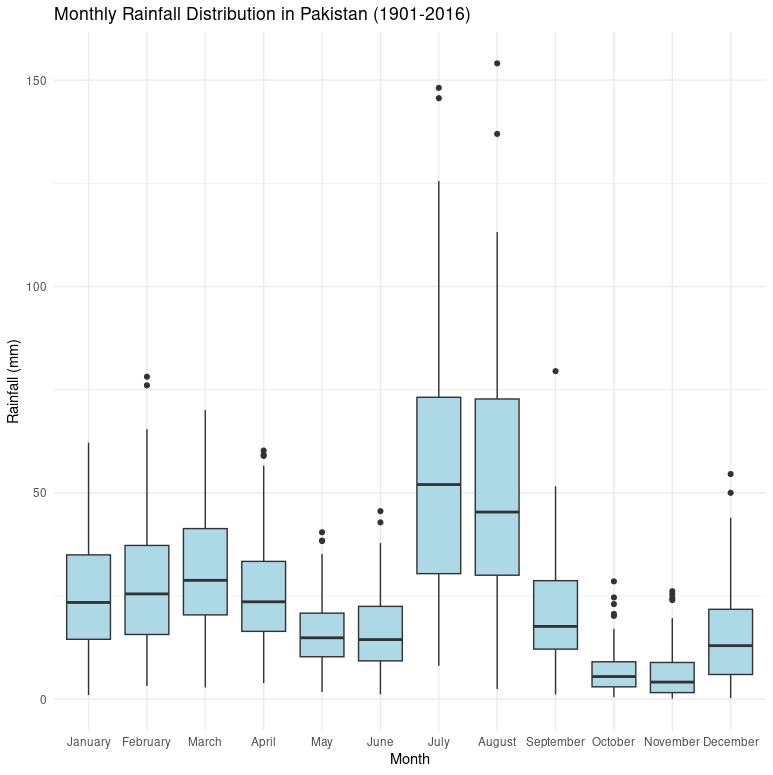
ggplot(rainfall\_data, aes(x = Month, y = Rainfall....MM.)) +

geom\_boxplot(fill = "lightblue") +

labs(title = "Monthly Rainfall Distribution in Pakistan (1901-2016)",

x = "Month",

y = "Rainfall (mm)") +

 theme\_minimal()

This boxplot provides insights into the distribution and variability of rainfall for each month, showing medians, quartiles, and potential outliers.

### Average Monthly Rainfall:

Finally, we analyze the average rainfall for each month:

# Aggregate data by month

monthly\_rainfall <- rainfall\_data %>%

group\_by(Month) %>%

summarise(Average\_Rainfall = mean(Rainfall....MM.))

# Plot average monthly rainfall

ggplot(monthly\_rainfall, aes(x = Month, y = Average\_Rainfall)) +

geom\_bar(stat = "identity", fill = "#808080") +

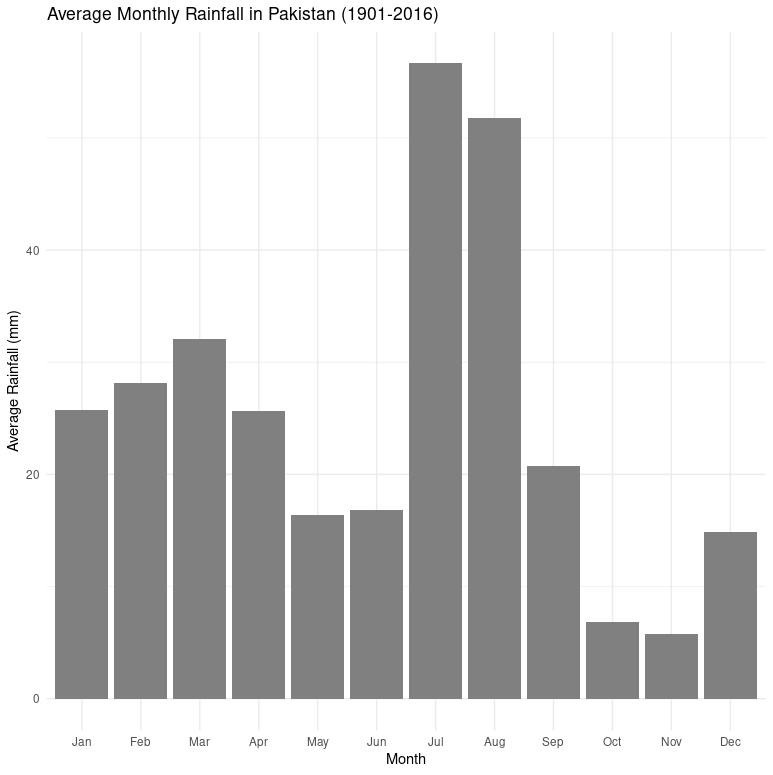
labs(title = "Average Monthly Rainfall in Pakistan (1901-2016)",

x = "Month",

y = "Average Rainfall (mm)") +

theme\_minimal() +

scale\_x\_discrete(labels = month.abb)



 Average rainfall for each month is calculated.

 A bar plot shows the average monthly rainfall, providing a clear view of seasonal patterns.

## Conclusion:

The visualizations created from the rainfall data provide a comprehensive view of the rainfall patterns in Pakistan from 1901 to 2016. Monthly and annual trends, as well as seasonal variations, are clearly depicted, helping to understand the historical rainfall behavior in the region.